

Research Article

Gender and Lexical Variation in Pitch, Duration, and Intensity: A Phonetic Study of Indonesian-Origin Words

¹Dedy Suhery, *¹Zahratul Idami, ²Rahmadsyah Rangkuti, ²Nina Tiya Hara, ²Ana Rahmawati Rangkuti

¹IAIN Langsa, Indonesia

²Universitas Sumatera Utara, Indonesia

*Corresponding author: zahra@iainlangsa.ac.id

Submitted: 13/08/2025

Revised: 12/10/2025

Accepted: 01/11/2025

How to cite this article: Suhery, D. Idami, Z., Rangkuti, R., Hara, N. T., & Rangkuti, A. R. (2025). Gender and lexical variation in pitch, duration, and intensity: A phonetic study of Indonesian-origin words. *IJELR: International Journal of Education, Language, and Religion*, 7(2), 155-164.
<https://doi.org/10.35308/ijelr.v7.i2.13165>

Abstract

This study investigates how gender influences suprasegmental features—specifically pitch, duration, and intensity—in the pronunciation of Indonesian-origin words. The research addresses the problem of limited phonetic studies examining gender-based acoustic variation in Indonesian, particularly in devoiced stop consonants. The objective is to analyze whether and how gender differences manifest in these prosodic elements. Using a descriptive qualitative method, data were collected from ten participants—five males and five females—from diverse ethnic backgrounds (Bataknese, Javanese, Karonese, and Malayunese), all residing in Medan, North Sumatra. Nine commonly used Indonesian-origin words were selected, and acoustic features were analyzed using PRAAT software. The results show that female speakers consistently produced higher pitch values across all words, with peaks reaching up to 499 Hz, while male speakers demonstrated lower and narrower pitch ranges, as low as 94 Hz. Duration values also varied more among female speakers, ranging from 0.097 ms to 0.469 ms, indicating more dynamic articulation patterns, although some male speakers also showed extended durations in specific contexts. Intensity levels ranged from 23 dB to 54 dB, with female speakers exhibiting greater variability in loudness—from soft to very loud—whereas male speakers maintained a more stable and moderate intensity. These findings suggest that pitch differences are primarily influenced by physiological factors such as vocal fold structure, while duration and intensity are more reflective of individual articulation style, emotional expressiveness, and speech clarity. Despite limitations in audio editing features in PRAAT, the study provides new insights into how gender and speaker identity shape suprasegmental variation in Indonesian phonetics.

Keywords

duration; gender variation; Indonesian phonetics; intensity; pitch



Introduction

In spoken language, the phonetic features of pitch, duration, and intensity play crucial roles in shaping the acoustic and perceptual characteristics of speech sounds (Steffman & Jun, 2019). These suprasegmental features not only influence how words are pronounced but also how they are perceived and understood by listeners. One particularly interesting phenomenon in this regard is devoicing, where voiced sounds lose their vocal fold vibration and are produced as voiceless (Hara et al., 2024). While devoicing has been widely studied in languages like Japanese, German, and English, its occurrence and phonetic realization in Indonesian-origin words remain relatively underexplored.

Indonesian, as an Austronesian language, exhibits a relatively simple phonemic inventory and prosodic system compared to stress-timed languages (Athanasopoulou et al., 2021). Nonetheless, in certain speech contexts—such as rapid speech, loanword adaptation, or dialectal influence—devoicing of typically voiced segments may occur. This raises questions about how such devoicing affects the pitch contour, segmental duration, and intensity levels of the affected words.

Several studies have been conducted regarding this issue. One of them revealed that the central focus in examining language through the lens of phonology is the investigation of phonological change or phonological processes—essential components that involve how sounds are structured and altered within a language system. A phonological process occurs when a speaker modifies a word's pronunciation by changing, adding, or inserting sounds at the beginning, middle, or end of the word (Diani & Azwandi, 2021). Such processes vary among different languages, dialects, and speakers, and they may evolve over time as languages respond to shifting social and linguistic contexts. Exploring these processes sheds light on the inner workings of language systems, how they are learned, and the influence of social and cultural dynamics in shaping them.

Using a constraint approach of Optimality Theory, Mose (2021) found that similar phonological processes occur in cases of borrowing between languages, whether related or unrelated; however, these studies primarily focused on structural or theoretical phonological aspects rather than empirical acoustic differences. Wu (2024), on the other hand, investigated pitch dynamism between male and female Thai students and found significant differences in pitch variance across genders. Similarly, other researchers have examined the phonological processes between male and female speakers of different languages, such as English–French (Pépiot & Arnold, 2020), English–Pakistani (Safeer et al., 2024), and Javanese (Mawarni et al., 2024).

Despite these contributions, previous studies tend to focus either on specific languages or theoretical frameworks and rarely integrate acoustic parameters such as pitch, duration, and intensity within the context of Indonesian-origin words. Moreover, gender-related variation in phonetic realization among speakers of borrowed or Indonesian-derived lexicons has received limited empirical attention. Therefore, this study seeks to fill this gap by providing a detailed phonetic analysis of pitch, duration, and intensity variations across male and female speakers when pronouncing Indonesian-origin words, thereby contributing new insights into the intersection of gender, phonetics, and lexical variation in the Indonesian linguistic context.

There are numerous types of phonological processes, such as aspiration, assimilation, insertion, deletion, voicing, devoicing, and nasalization. Among these, devoicing is a key focus in this study. This is because Indonesian exhibits a phonological pattern known as final-obstruent devoicing (Suherly et al., 2023), where voiced consonants like /b/, /d/, /g/, /z/, and /ʒ/ become voiceless counterparts /p/, /t/, /k/, /s/, and /ʃ/ at the end of words. In practice, this means that in Indonesian, word-final consonants often lose their voicing. For example, *murid* (student) is pronounced [murit], *abad* (century) becomes [abat], and *ajaib* (magic) is pronounced [ajaip]. Several explanations may



account for this phenomenon. Iwasaki et al. (2022) say that one is articulatory ease—voiceless sounds are generally easier to produce at word-final positions because voicing requires greater airflow and muscular control. Another reason may be the effort to enhance perceptual contrast between adjacent sounds, particularly vowels, which are typically voiced. Additionally, this process might be influenced by the prosodic features of the language, such as its stress patterns, intonation, or rhythm. The current study seeks to investigate the acoustic properties of devoiced pronunciation in a selected set of Indonesian-origin words. By analyzing changes in pitch, duration, and intensity, this research aims to identify consistent phonetic patterns associated with devoicing. In other words, the study examines the acoustic consequences of devoicing in Indonesian-origin words, with a particular focus on three key phonetic parameters: pitch contour, segmental duration, and intensity. It explores how the process of devoicing influences the pitch trajectory of these words and examines the extent to which the temporal properties of speech segments differ between devoiced and fully voiced pronunciations. These concerns emphasized a major gap that distinguishes the present research from previous studies, which often focused on theoretical phonology or limited cross-gender comparisons without systematically exploring acoustic parameters in Indonesian-origin lexicons.

In addition, the study analyzes how vocal intensity shifts when comparing devoiced forms to their normally voiced counterparts. A further objective is to determine whether there are consistent and identifiable acoustic cues—across pitch, duration, and intensity—that reliably signal devoicing across different speakers of Indonesian. Furthermore, the study contributes to a better understanding of how these acoustic cues interact in shaping the phonological structure of Indonesian words when devoicing occurs.

Ultimately, this study not only adds to the body of knowledge in phonetic and phonological research within the Indonesian linguistic context but also offers insights into broader cross-linguistic patterns of voicing and its acoustic correlates. Based on these aims, the present research is guided by the following research questions: how does the process of devoicing in Indonesian-origin words affect acoustic properties such as pitch, duration, and intensity?; are there consistent and identifiable acoustic cues that distinguish devoiced pronunciations from their fully voiced counterparts across different speakers?; and to what extent do gender differences influence the acoustic realization of devoicing in Indonesian-origin words?

Method

In order to clearly and completely depict social or human issues—with a focus on phonological issues—descriptive qualitative research was used in this study. Creswell (2014) defines qualitative research as a technique for examining and comprehending the significance of individuals or groups in connection to societal issues. It can be applied to more thoroughly interpret, research, or understand a certain aspect of human attitudes, beliefs, or behavior. This study focused on carefully analyzing and describing a specific phenomenon and situation. In this case, material was gathered, examined, and assessed by the researchers before being described. This method aims to address every aspect of the research problem. When the researchers examined the phenomena of stop consonants devoicing in end words using both optimality and generative phonology theory.

The study involved a total of ten participants, evenly split between male and female speakers, purposively. These individuals come from diverse ethnic backgrounds, including Batakese, Javanese, Karonese, and Melayunese, and all reside in Medan, North Sumatra. To ensure age diversity, participants were chosen to represent five distinct age brackets: 10–20, 21–30, 31–40, 41–50, and 51–60 years. Each age group includes one male and one female speaker. For instance, the youngest group consists of a 15-year-old male (M15) and a 19-year-old female (F19), while the oldest



group includes a 55-year-old male (M55) and a 53-year-old female (F53). This balanced distribution enables the analysis of age- and gender-related patterns in the phonetic data.

This research specifically concentrated on three voiced stop consonants: the bilabial /b/, alveolar /d/, and velar /g/. Participants were instructed to pronounce a set of selected Indonesian words that contained these target consonants. The chosen words— *akad* (agreement), *jilid* (binding), *wujud* (form or shape), *kutub* (pole), *sebab* (cause), *tertib* (orderly), *caleg* (legislative candidate), *dialog* (dialogue), and *warteg* (local food stall)—were selected for their phonetic relevance and lexical familiarity. Each participant's pronunciation was recorded using a smartphone device, a TASCAM DR-05X stereo hand-held digital audio recorder, to ensure consistent data collection across all speakers. These items served as the primary data for analyzing the acoustic properties associated with the target consonants.

Following data collection, PRAAT software was used for transcription and analysis. In PRAAT, pitch analysis relies heavily on configurable parameters like pitch range, analysis method, and time step, which determine how accurately and meaningfully pitch contours are extracted and displayed. According to Ladefoged & Johnson (2010), PRAAT is a reliable tool for speech sound analysis because it allows users to explore a wide range of processes, including creating spectrograms and pitch analyses, examining how the ear interprets sounds, synthesizing speech in articulatory terms, utilizing neural nets, using optimality theory to describe phonetic events, and much more. The vocalizations were measured using PRAAT, which also provided each participant's pitch (dB), frequency (Hz), and duration (ms) data. Shang (2016) mentions that the use of PRAAT shows a great influence on voice.

Results

The results of the study indicate that devoicing was a particular type of pronunciation error among all the phonemes that the participants produced. The research findings were further displayed in the chart that follows.

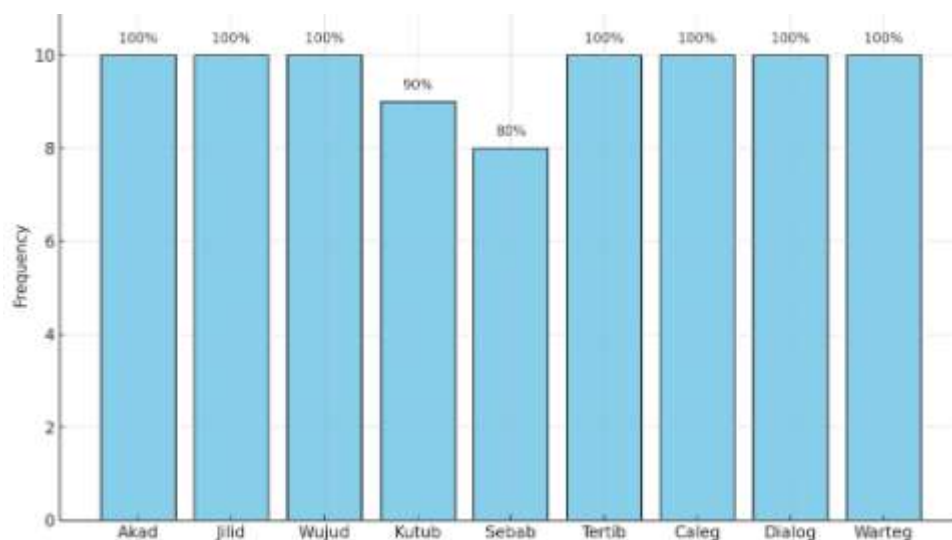


Figure 1. Devoicing produced by participants

Based on the above chart, the researchers discovered that only two participants, F21 and M47, were able to pronounce the words [kutub] and [səbab] without any devoicing attempts, while nearly all participants produced the stop sounds [b], [d], [g], which experienced devoicing to create the sounds [p], [t], [k].

Pitch

The high and low levels of speech flow define pitch, a kind of suprasegmental feature (Choi, S., & Kang, 2023). Based on the frequency of the sound waves, pitch is the sense of a sound's height or low volume. Tone employs pitch to determine the meaning of words or phrases, whereas intonation uses pitch to express different emotions, attitudes, or functions in utterances. The researchers show the pitch measurement results to each participant in the table below.

Table 1. Pitch of each participant according to PRAAT

No	Participant	[a.kad]	[ji.lid]	[wu.jud]	[ku.tub]	[se.bab]	[ter.tib]	[ca.leg]	[dia.log]	[war.teg]
1	F19	206 hz	144 hz	340 hz	171 hz	268 hz	179 hz	281 hz	247 hz	178 hz
2	M15	106 hz	111 hz	111 hz	114 hz	108 hz	98 hz	94 hz	107 hz	109 hz
3	F21	264 hz	499 hz	226 hz	228 hz	244 hz	231 hz	224 hz	205 hz	321 hz
4	M24	165 hz	242 hz	304 hz	353 hz	165 hz	298 hz	361 hz	171 hz	278 hz
5	F36	194 hz	392 hz	321 hz	358 hz	184 hz	421 hz	201 hz	142 hz	153 hz
6	M31	183 hz	440 hz	472 hz	373 hz	126 hz	456 hz	124 hz	122 hz	126 hz
7	F45	178 hz	179 hz	238 hz	271 hz	270 hz	217 hz	255 hz	233 hz	159 hz
8	M47	134 hz	138 hz	137 hz	173 hz	127 hz	183 hz	140 hz	140 hz	121 Hz
9	F53	391 hz	473 hz	224 hz	463 hz	448 hz	426 hz	223 hz	198 hz	432 hz
10	M55	218 hz	202 hz	230 hz	202 hz	222 hz	199 hz	208 hz	163 hz	373 hz

The table indicates that female participants generally exhibit higher pitch values across all words. For example, participant F21 produced pitch values of 499 Hz for [ji.lid], 321 Hz for [wu.jud], and 244 Hz for [Se.bab]. In contrast, participant F53 consistently produces extremely high pitches of 473 Hz for [ji.lid], 448 Hz for [Se.bab], and 432 Hz for [war.teg]. Male participants, on the other hand, consistently generate lower pitch values, often below 200 Hz. For instance, participant M15 has uniform pitch values ranging from 94 to 114 Hz across all words, while participant M47 shows a similar low range of 134 to 140 Hz.

These findings reflect the typical physiological differences in pitch between male and female speakers, which are influenced by vocal fold length and mass. Notably, the lowest pitch recorded was 94 Hz, produced by M15 while saying the word [caleg], whereas the highest pitch recorded was 499 Hz, produced by F21 for the word [jilid].

Duration

Duration is another important factor in sound differentiation. Time and the temporal characteristics that characterize speech sounds, such as words, paragraphs, and syllables, are linked to all phonetic energy in acoustics (Irawan, 2017). Additionally, duration reveals if a language employs a long (geminate) consonant, a short consonant, or a short vowel. Gósy (2023) comes to the conclusion that quiet and pauses are related to duration as well.

All participants pronounce devoiced stop consonants at different lengths, according to the results in the table below. The time measurements of the participants' pronunciation of the provided words are displayed in the table. F21 took the shortest time, taking 0.097506 ms to pronounce the letter [p] in the word [kutup], while F36 took the longest, taking 0,469728 ms to pronounce the letter [t] in the word [jilit].

Table 2. Duration of each participant according to PRAAT

No	Participants	[a.kad]	[ji.lid]	[wu.ju]	[ku.tub]	[se.bab]	[ter.tib]	[ca.leg]	[dia.log]	[war.teg]
1	F19	0.2078 1 ms	0.2058 05 ms	0.1893 88 ms	0.1849 43 ms	0.1717 01 ms	0.2130 61 ms	0.2128 80 ms	0.1952 83 ms	0.1902 04 ms
2	M1 5	0.3199 77 ms	0.2436 48 ms	0.2403 63 ms	0.2698 19 ms	0.3120 18 ms	0.2218 59 ms	0.2713 61 ms	0.2356 46 ms	0.3038 55 ms
3	F21	0.1508 39 ms	0.1875 74 ms	0.1058 50 ms	0.0975 06 ms	0.1233 47 ms	0.1150 01 ms	0.2041 71 ms	0.2655 78 ms	0.2042 63 ms
4	M2 4	0.2260 23 ms	0.2652 15 ms	0.2267 57 ms	0.1438 55 ms	0.1902 95 ms	0.1758 73 ms	0.2445 35 ms	0.2045 35 ms	0.1035 83 ms
5	F36	0,2723 34 Ms	0,4697 28 ms	0.1435 83 ms	0,3945 80 ms	0.1792 29 ms	0.1861 32 ms	0.1988 21 ms	0.1822 00 ms	0.2718 59 ms
6	M3 1	0.1433 11 ms	0.1273 48 ms	0.2634 92 ms	0.2306 28 ms	0.1070 29 ms	0.1527 44 ms	0.1818 59 ms	0.1821 32 ms	0.1687 07 ms
7	F45	0,3010 66 ms	0,2750 34 Ms	0,4392 06 Ms	0,3937 41 ms	0,3945 80 ms	0,4197 28 ms	0,3845 80 ms	0,3812 70 ms	0,3839 00 ms
8	M4 7	0.1979 14 ms	0.1922 00 ms	0.1439 46 ms	0.2821 777 ms	0.1976 42 ms	0.2821 77 ms	0.1424 94 ms	0.1025 85 ms	0.1792 29 ms
9	F53	0.2443 54ms	0.1297 96 ms	0.1554 65 ms	0.3010 43 ms	0.3209 07 ms	0.1988 21 ms	0.2446 26 ms	0.2019 05 ms	0.2319 27 ms
10	M5 5	0.1269 54 ms	0.2513 38 ms	0.3059 41 ms	0.2830 84 ms	0.1277 70 ms	0.3980 05 ms	0.2976 87 ms	0.2712 93 ms	0.3037 64 ms

Female speakers tend to exhibit a wider variation in duration, with both shorter and longer values. For instance, F21 demonstrates very short durations, such as [ku.tub], which is 0.097 ms. In contrast, F45 consistently produces longer durations, ranging from 0.393 to 0.439 ms for many words. Meanwhile, male speakers, such as M15 and M55, also produce longer durations. M15 has notable durations like [a.kad] at 0.319 ms and [se.bab] at 0.312 ms. Similarly, M55 produces [wu.ju] at 0.305 ms and [ter.tib] at 0.398 ms.

Intensity

Martin (2021) state that intensity is proportional to the amplitude, or size of displacement, in a sound vibration, and is measured in decibels (dB). To put it another way, intensity is a loud sound that reflects the loudness or softness of the sound. According to Fry and Lehiste, as cited in Irawan (2017), loud noises are influenced by various acoustic phonetic components. The intensity data for each participant is shown in the table below, with the lowest being 23 dB for F53 when producing the word [warteg] and the highest reaching 54 dB for F45 when uttering the word [akad].

Table 3. Intensity of each participant according to PRAAT

No	Participant	[a.kad]	[ji.lid]	[wu.ju]	[ku.tub]	[se.bab]	[ter.tib]	[ca.leg]	[dia.log]	[war.teg]
1	F19	44 dB	44 dB	44 dB	41 dB	45 dB	38 dB	44 dB	46 dB	42 dB
2	M15	43 dB	45 dB	38dB	37 dB	30 dB	41 dB	41 dB	36 dB	32 dB
3	F21	38 dB	36 dB	33 dB	37 db	43 dB	33 dB	34 dB	39 dB	32 dB
4	M24	39 db	42 dB	39 dB	33 dB	37 dB	39 dB	40 dB	41 dB	45 dB
5	F36	43 dB	42 dB	32 dB	31 dB	45 dB	31 dB	38 dB	50 dB	38 dB
6	M31	44 db	32 dB	39 dB	32 dB	46 dB	35 dB	42 dB	41 dB	41 dB
7	F45	54 dB	49 dB	42 dB	44 dB	51 dB	37 dB	36 dB	36 dB	52 dB
8	M47	41 dB	37 dB	37 dB	41 db	41 dB	41 dB	44 dB	43 dB	42 dB
9	F53	42 dB	26 dB	34 dB	28 db	29 dB	28 dB	26 DB	40 dB	23 dB
10	M55	44 dB	40 dB	33 dB	39 dB	53 dB	36 dB	31 db	36 dB	32 dB



The measured speech intensities for the nine Indonesian-origin words span a range from 23 dB to 54 dB, which is a typical variation in conversational loudness. The highest Intensity Values F45 stands out clearly as the loudest speaker: 54 dB on [a.kad], 52 dB on [war.teg], and it maintains high intensity across most words, often in the upper 40s and low 50s, M55 shows localized high intensity: 53 dB on [se.bab] — this could suggest emphasis or stress on that word specifically. These high dB values may indicate stronger vocal projection, expressive emphasis, or possibly proximity to the microphone during recording.

Then the lowest Intensity Values F53 is the quietest speaker in the dataset: 23 dB on [war.teg], 26 dB on [ji.lid]. Several other values are below 30 dB (e.g., [ku.tub] = 28 dB, [ter.tib] = 28 dB). This suggests a consistently soft speech style, which could be due to personality traits, speech habits, physical condition (e.g., older speakers), or technical factors like microphone distance. In addition, from Gender-Based Patterns, we can state that female speakers show greater variation in intensity, ranging from very soft (F53) to very loud (F45). F45 and F19 show consistently higher intensity (typically above 40 dB), and F45 in particular may indicate emphatic or expressive speech. F21 and F53 fall on the quieter end, particularly F53. This suggests that female speakers in this sample are more variable, possibly due to personal speaking styles, confidence, or age differences.

Meanwhile, male speakers have a more consistent intensity range, generally between 36–44 dB. M31 and M47 show balanced, moderate loudness across all words, averaging around 41–42 dB. M15 and M55 dip lower on some words: M15: 30 dB on [sə.bab], 32 dB on [war.tæg], M55: 31 dB on [ca.leg], 32 dB on [war.teg]. These male speakers do not reach the high peaks of F45 but also avoid the deep troughs of F53.

Henceforth, the results demonstrated that both male and female speakers possessed phonetic diversity, supporting the research findings of Traunmüller & Eriksson (2000) on the duration of vowel and consonant sounds produced by men, women, and children.

Discussion

The acoustic analysis of the selected Indonesian-origin words reveals notable gender-based variations in pitch, duration, and intensity, offering insights into phonetic diversity shaped by physiological, sociolinguistic, and contextual factors.

Devoicing Patterns

Only two participants, F21 and M47, consistently avoided devoicing, producing [kutub] and [səbab] with the original voiced stops [b], [d], and [g]. The majority demonstrated a tendency to devoice these consonants, transforming them into [p], [t], and [k]. This pattern suggests a phonological shift or variation in articulatory precision, possibly influenced by speech style, effort, or regional tendencies. This result is in line with Suhery et al., (2023)'s research on phonological patterns regarding final-obstruent devoicing as a type of phonological process.

Pitch Variation

Pitch measurements confirm a clear gender distinction; female speakers tend to produce higher pitch values, consistent with established physiological traits such as shorter and thinner vocal folds. F21 and F53 are notable for extreme pitch heights—F21 reached 499 Hz on [ji.lid], and F53 consistently surpassed 400 Hz on multiple words. Meanwhile, male speakers maintained lower and more stable pitch ranges, frequently below 200 Hz (e.g., M15 between 94–114 Hz, M47 around 134–140 Hz). This reflects typical sexual dimorphism in phonation and supports findings by Choi & Kang (2023) on suprasegmental features in speech. These findings are supported by Wu (2024)'s study on phonological issues of Thai students.



Duration Patterns

Speech duration presents a wider variability among female speakers: F21 produced some of the shortest durations (e.g., 0.097 ms for [kutub]), indicating brisk articulation. F45 and F36, on the other hand, consistently displayed longer durations, potentially signifying deliberate speech or expressive elongation.

Male speakers such as M15 and M55 also produced extended durations on several words, suggesting individual differences beyond gender. These findings align with Gósy (2023)'s assertion that temporal speech features are influenced by pauses, rhythm, and prosodic structuring.

Intensity Differences

Intensity results further highlight gender-based contrasts: F45 emerged as the most vocally expressive speaker with peak intensity at 54 dB ([a.kad]) and consistently loud delivery. F53, in contrast, was the quietest (23 dB on [war.təg]), possibly reflecting age-related factors or microphone distance. Female speakers exhibited greater variation in vocal intensity, suggesting differing degrees of confidence, emotional expression, or speech style. Male speakers maintained a narrower range (generally 36–44 dB), with moderate loudness and fewer peaks or troughs. This acoustic variation corroborates Martin (2021) and Irawan (2017), underscoring intensity as a product of vocal effort, amplitude, and phonetic components.

Gender and Phonetic Diversity

The observed contrasts in pitch, duration, and intensity affirm the presence of phonetic diversity among both male and female speakers, supporting the work of Traunmüller & Eriksson (2000) on age and gender-related vocal attributes. In the end, the female speakers show broader variability and more extremes. In comparison, male speakers tend toward stability and moderation in delivery. These patterns are influenced by anatomical factors, individual speaking habits, sociocultural roles, and technical constraints during data collection.

Conclusion

Taken together, the results reveal that pitch, duration, and intensity patterns among Indonesian speakers are influenced by a complex interplay of biological, linguistic, and sociocultural factors. The consistently higher pitch values produced by female speakers align with well-documented physiological explanations related to vocal fold size and mass. However, the degree of pitch variation observed—especially among speakers such as F21 and F53—suggests that biological determinants alone cannot fully account for the differences. Social factors such as expressive style, gender identity performance, and speech accommodation may also shape how speakers modulate their voice to convey affect, politeness, or emphasis within Indonesian discourse contexts.

In terms of duration, the variability found across both male and female speakers indicates that temporal features are not merely reflections of gender-based physiological traits but are instead tied to individual articulatory strategies and communicative intent. Longer stop durations among certain male speakers, for instance, may reflect careful articulation or emphasis patterns shaped by pragmatic or stylistic preferences rather than by biological constraints.

Intensity differences further demonstrate that female speakers exhibit greater dynamic range in loudness, which may correspond to social expressiveness or interactional norms in Indonesian communication. The broader range of intensity—from soft to strong delivery—points to how emotional stance and self-presentation intersect with acoustic realization. Male speakers' relatively stable intensity levels, meanwhile, suggest a preference for moderate prosodic control, which could reflect cultural expectations surrounding male speech behavior.



Overall, it can be concluded that gender-based physiological differences clearly influence pitch, with female speakers exhibiting higher frequency values. Duration appears more closely tied to individual articulation patterns than strictly to gender, as both male and female speakers showed variability. Intensity displayed greater variation among female speakers, particularly between F45 and F53, suggesting influences from speaking style or emotional expressiveness.

Even though the PRAAT software's inability to be edited is one of its intrinsic limitations, and the audio file was trimmed using the WavePad sound editing program, the findings are important for understanding phonetic variation in devoiced stops among Indonesian speakers. They also highlight how gender, speaker identity, and individual style affect suprasegmental features like pitch, duration, and intensity in spoken language.

Acknowledgement

The authors would like to express their sincere gratitude to all participants who contributed their time and voices to this study. We also extend our appreciation to colleagues and peer reviewers for their valuable feedback throughout the research process. Special thanks go to the institutions and academic communities that supported the completion of this work, both technically and administratively. This study would not have been possible without the collaborative efforts and shared commitment of all four authors, whose contributions were equally significant to the design, execution, and writing of this article.

References

- Athanasopoulou et al. (2021). Prosodic prominence in a stressless language: An acoustic investigation of Indonesian. *Journal of Linguistics*, 57(4), 695–735.
<https://doi.org/doi:10.1017/S0022226721000141>
- Choi, S., & Kang, O. (2023). The roles of suprasegmental features in assessing paired speaking tasks in high-stakes language assessment. *System*, 119, 103183.
<https://doi.org/https://doi.org/10.1016/j.system.2023.103183>.
- Creswell, J. W. (2014). Research Design: Qualitative, Quantitative, and Mix Method Approach. In *University of Nebraska, Lincoln* (4th ed.). Sage.
- Diani & Azwandi. (2021). Phonological change processes of English and Indonesian language. *Journal of Applied Linguistics and Literature Vol 6 No 1*, 6(1), 133–148.
<https://doi.org/https://doi.org/10.33369/joall.v6i1.13642>
- Gósy, M. (2023). Occurrences and durations of filled pauses in relation to words and silent pauses in spontaneous speech. *Languages*, 8(1), 79.
<https://doi.org/https://doi.org/10.3390/languages8010079>.
- Hara et al. (2024). Devoicing of final voiced stop consonants in Indonesian. *Ethical Lingua: Journal of Language Teaching and Literature.*, 11(2).
<https://doi.org/https://doi.org/10.30605/25409190.704>
- Irawan, Y. (2017). *Fonetik akustik: Sebuah pengantar telaah wujud akustik bahasa*. CV Angkasa.
- Iwasaki et al. (2022). Retention of devoiced vowels in Tokyo Japanese: Evidence from lip articulation. *The Journal of the Acoustical Society of America*, 152(4).
<https://doi.org/https://doi.org/https://doi.org/10.1121/10.0016299>
- Lodefog & Johnson (2010). *A Course in phonetics sixth edition*. Cengage Learning.
- Martin, P. (2021). Fundamental frequency and intensity. In *Speech Acoustic Analysis*. Wiley.
<https://doi.org/https://doi.org/10.1002/9781119808411.ch7>.
- Mawarni et al. (2024). Analysis of Javanese language prosody (acoustic phonetic study). *Fonologi : Jurnal Ilmuan Bahasa Dan Sastra Inggris*, 2(4), 63–75.
<https://doi.org/https://doi.org/https://doi.org/10.61132/fonologi.v2i4.1136>
- Mose, E. (2021). Phonological processes in Ekegusii borrowing: A constraint-based approach. *JL3T*



- (*Journal of Linguistics, Literature and Language Teaching*), 7(2), 83–100.
<https://doi.org/https://doi.org/10.32505/jl3t.v7i2.3233>
- Pépiot & Arnold. (2020). Cross-gender differences in English/French bilingual speakers: A multiparametric study. *Perceptual and Motor Skills*, 128, 153–177.
<https://doi.org/https://doi.org/10.1177/0031512520973514>
- Safeer et al. (2024). Gender-based study of paired monophthongs: A sociophonetics approach. *3L The Southeast Asian Journal of English Language Studies.*, 30(2), 231–262.
<https://doi.org/https://doi.org/10.17576/31-2024-3002-15>
- Shang, C. Y. (2016). Research on the application of Praat in English pronunciation class. *Journal of Mudanjiang University*, 25(4), 4.
- Steffman & Jun. (2019). Perceptual integration of pitch and duration: Prosodic and psychoacoustic influences in speech perception. *The Journal of the Acoustical Society of America*, 146(3), EL251–EL257. <https://doi.org/https://doi.org/10.1121/1.5126107>
- Suhery et al. (2023). Interlingual interference in multilingual students: An optimality theory approach. *RETORIKA: Jurnal Ilmu Bahasa*, 9(3), 275–283.
<https://doi.org/https://doi.org/10.55637/jr.9.3.8875.275-283>
- Traunmüller & Eriksson. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America*, 107(6), 3438–3451.
<https://doi.org/https://doi.org/10.1121/1.429414>
- Wu, A. (2024). Gender and situational variation of pitch dynamism in Thai speakers. *NIDA Journal of Language and Communication*, 29(45), 60–80.

